



QUANTIFYING RISK, ENABLING OPPORTUNITY

# ASSA 2020 Data Science Survey: Feedback presentation

ASSA Wider Fields Committee  
ASSA  
Date: 8 December 2020

# Content

**Introduction**

**Aims of the survey**

**Members**

**Organization**

**Tools and techniques**

**Risks and opportunities**

**The survey**

# Introduction

**The Actuarial Society of South Africa has issued three surveys around data science in the actuarial community:**

- 2016 (233 respondents)
- 2018 (166 respondents)
- 2020 (312 respondents)

# Aims of the survey

- **Evaluate the inroads data science is making within the members of the actuarial profession and the organisations they work for.**
- **Consider the role and responsibility of the actuary in data science**
- **Approaches to data science in organisations**
- **Which techniques and tools are being used by actuaries within their daily work**
- **Highlight beliefs held by the actuarial profession about Data Science**
- **Biggest contributors to Data Science success and failure**

# Agenda

Members

Organizations

Tools and  
Techniques

Risks and  
opportunities

The survey

# Agenda

Members

Organizations

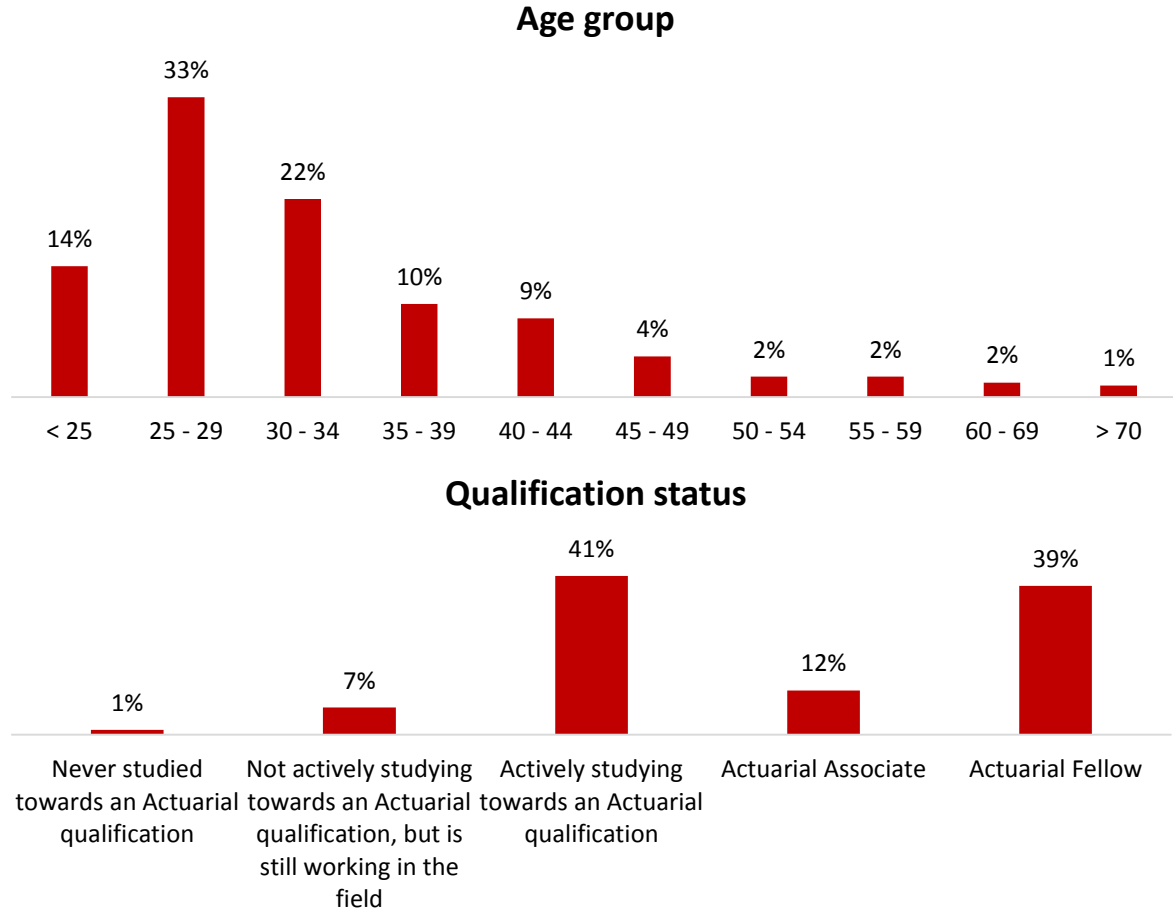
Tools and  
Techniques

Risks and  
opportunities

The survey

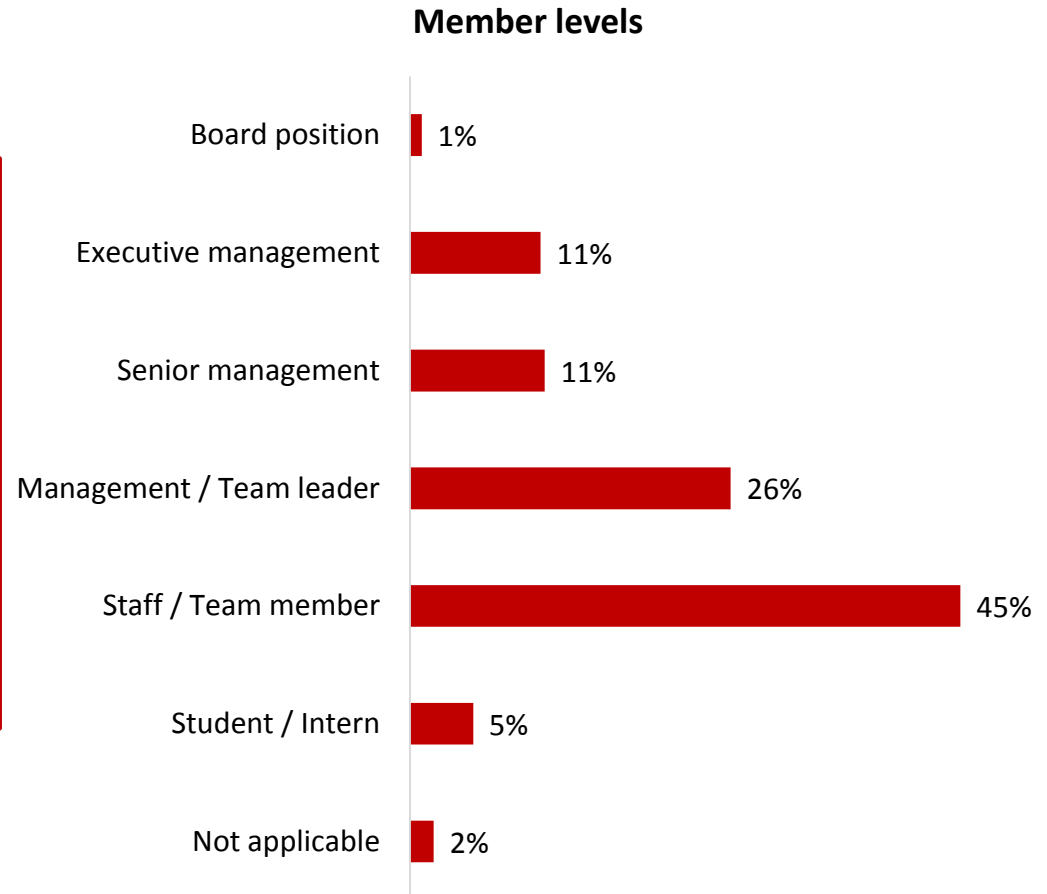
# Age & qualification status

Respondents were generally from younger age groups. There was representation across different qualification statuses.



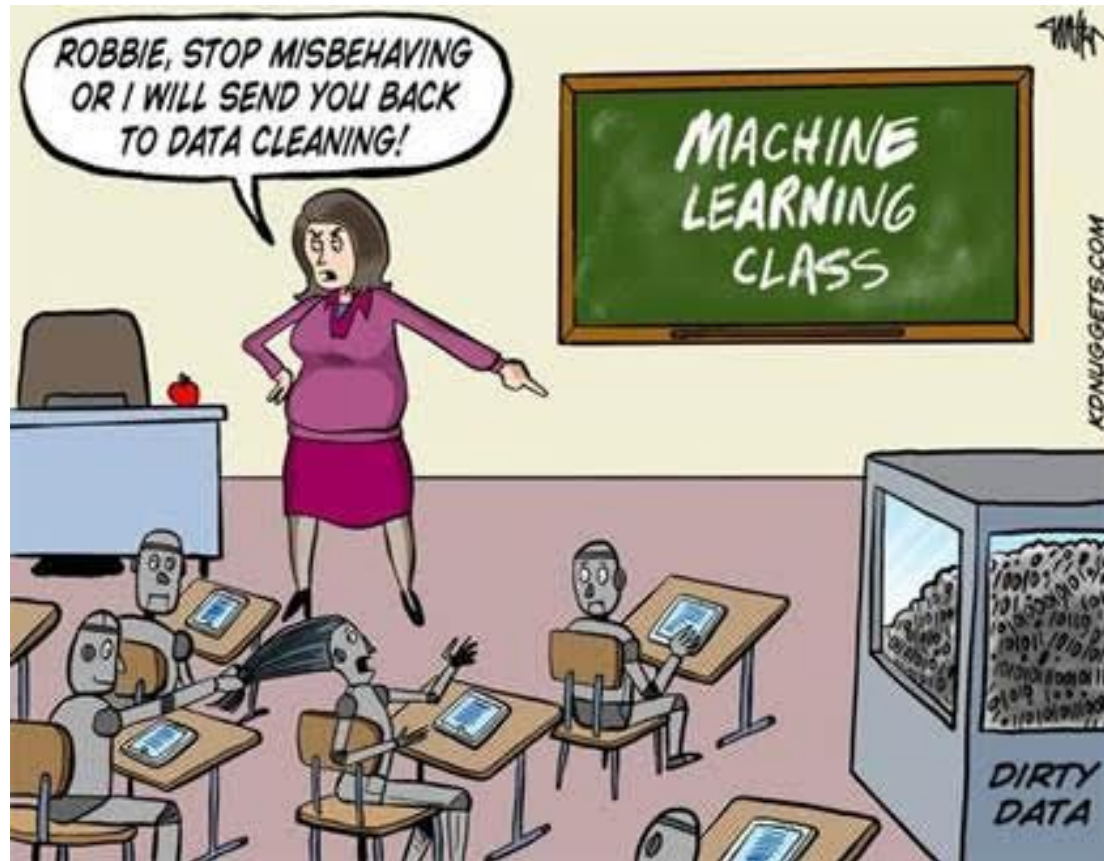
# Member levels

Representation was decent across the different member levels with the exception of board members and students being under-represented as expected.





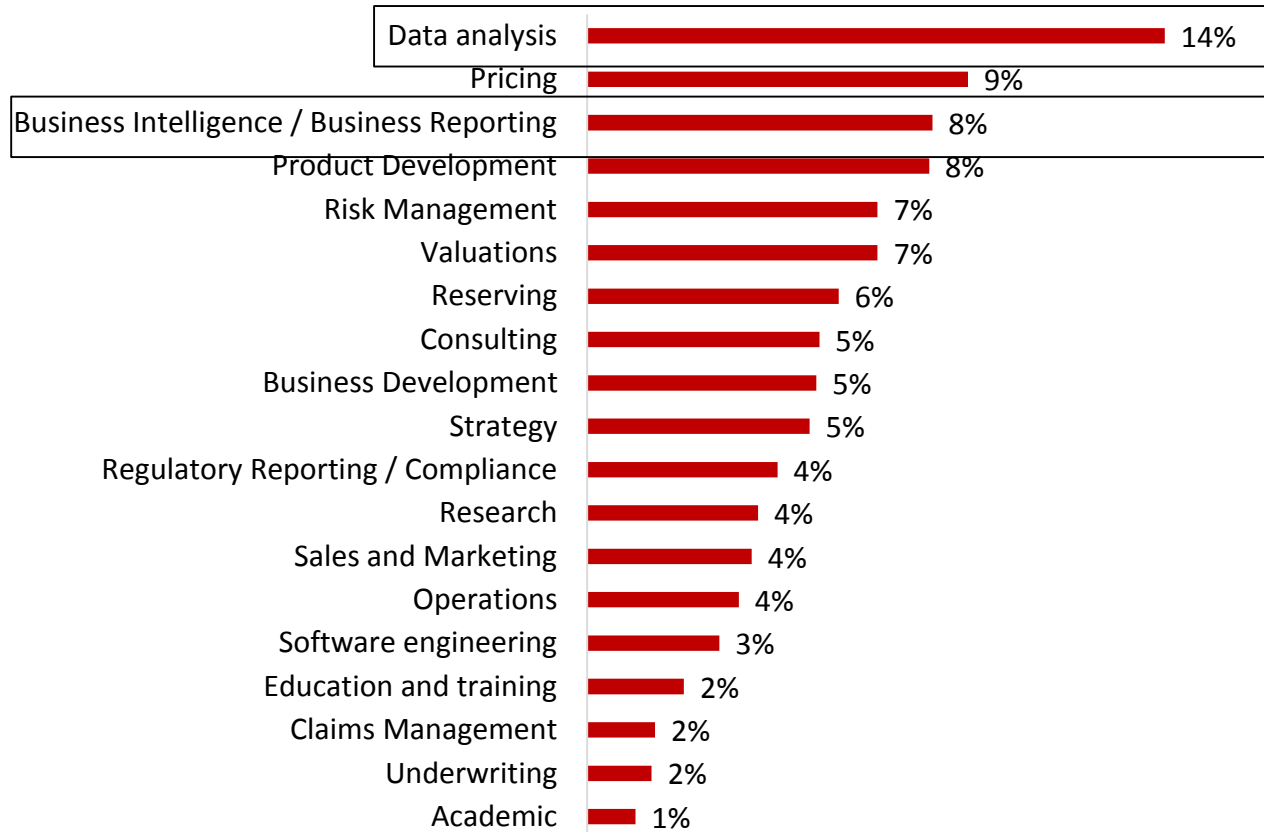
# Focus areas



# Focus areas

Data analysis and BI form a large part of both Data Science and Actuarial work

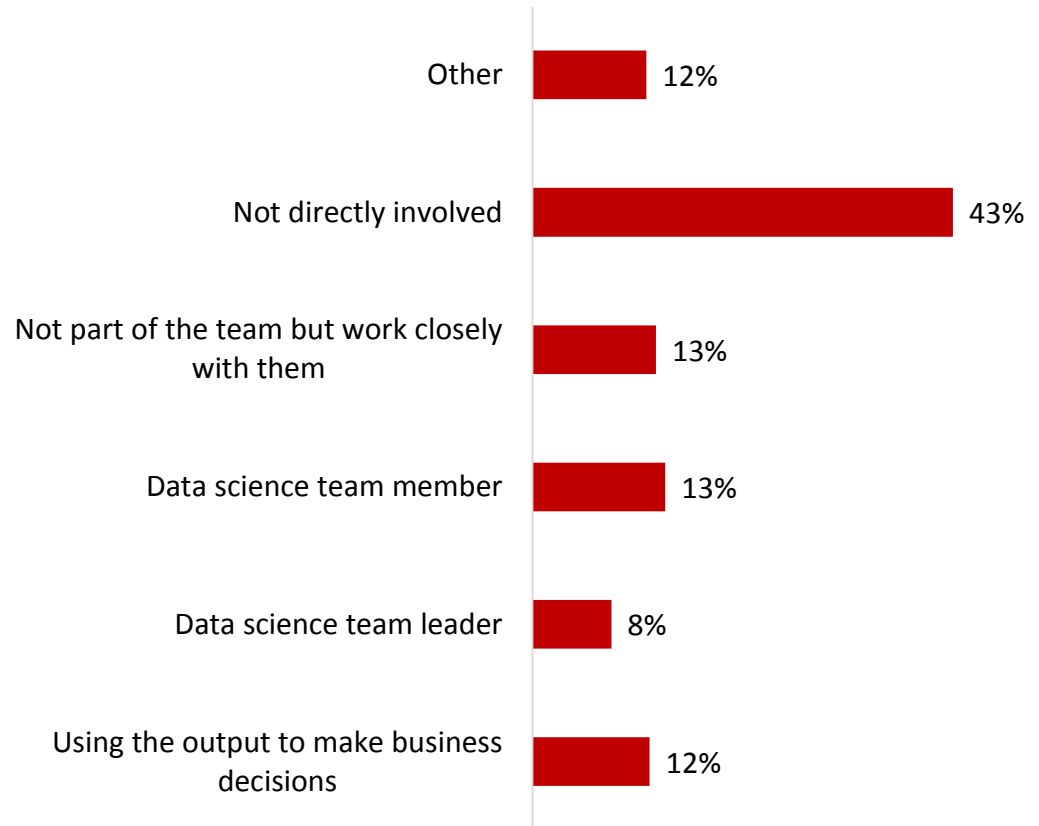
## Focus areas



# Involvement with data science team

**Involvement with data science team**

There is still a large portion of actuaries who are not directly involved with data science teams.

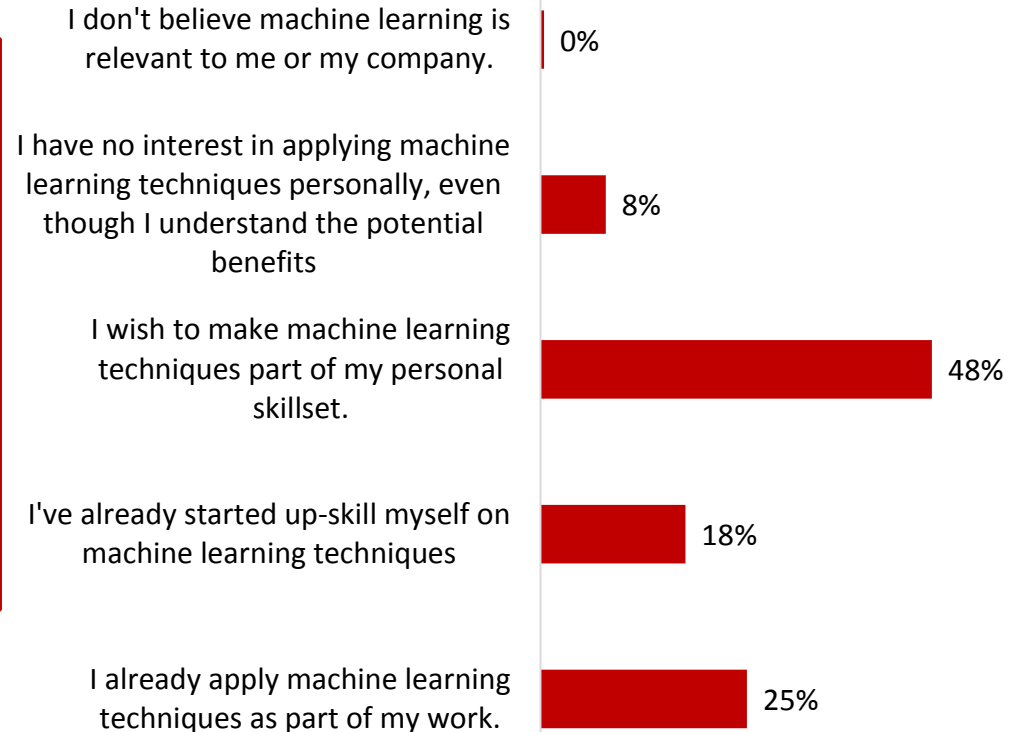


# Machine learning adoption

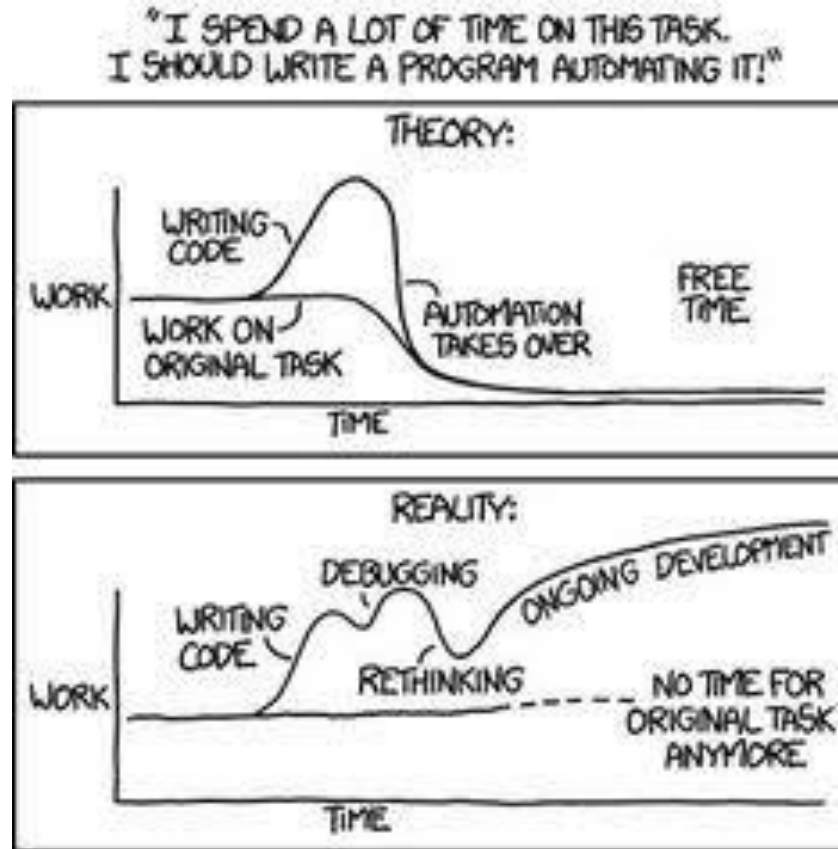
## Machine learning attitude

All respondents (barring one) believed machine learning is relevant to themselves or their companies.

92% of respondents expressed a need for being upskilled in machine learning.



# The actuary's role in model production

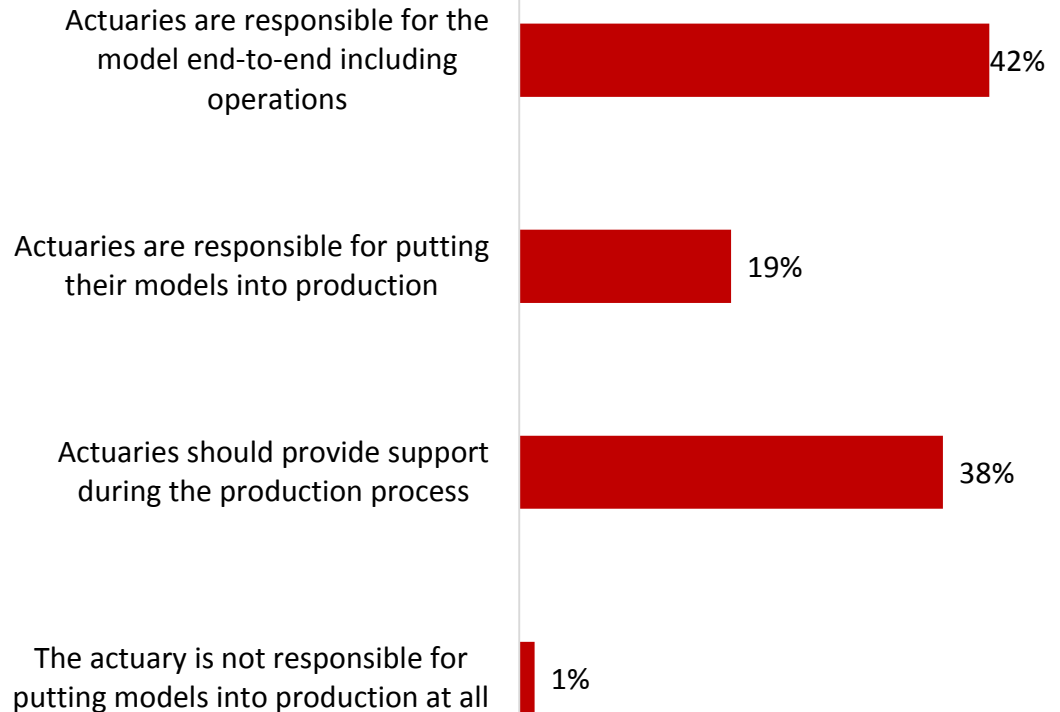


# Actuaries' role in model production

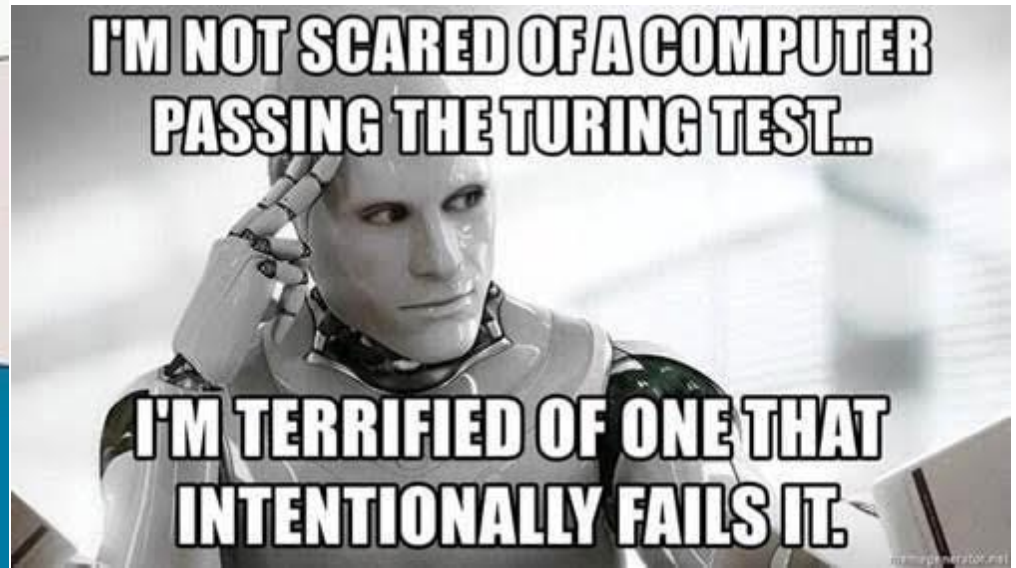
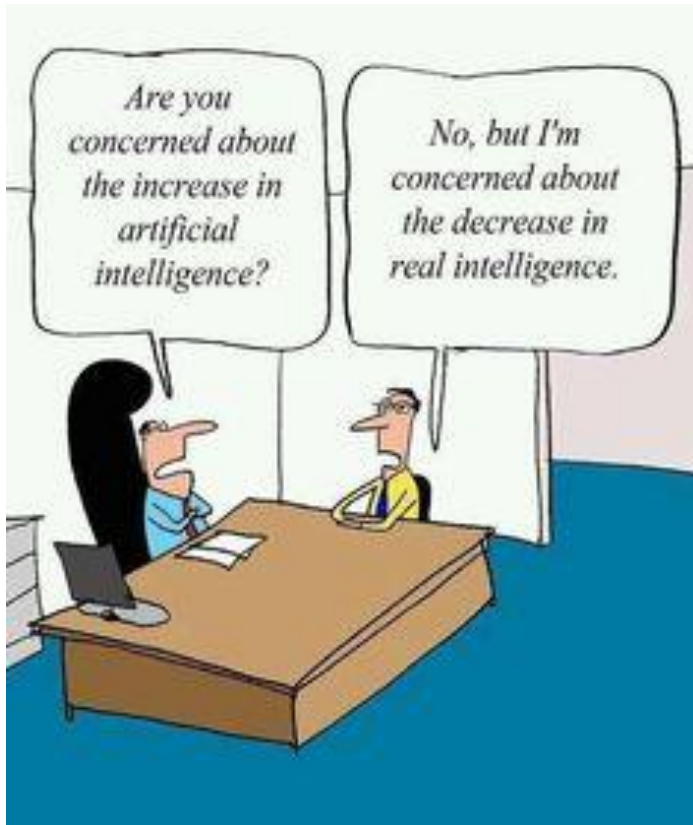
## Actuaries' role in model production

Almost all respondents felt that actuaries' roles extend into production.

The extent of the need for involvement seems to be distributed.

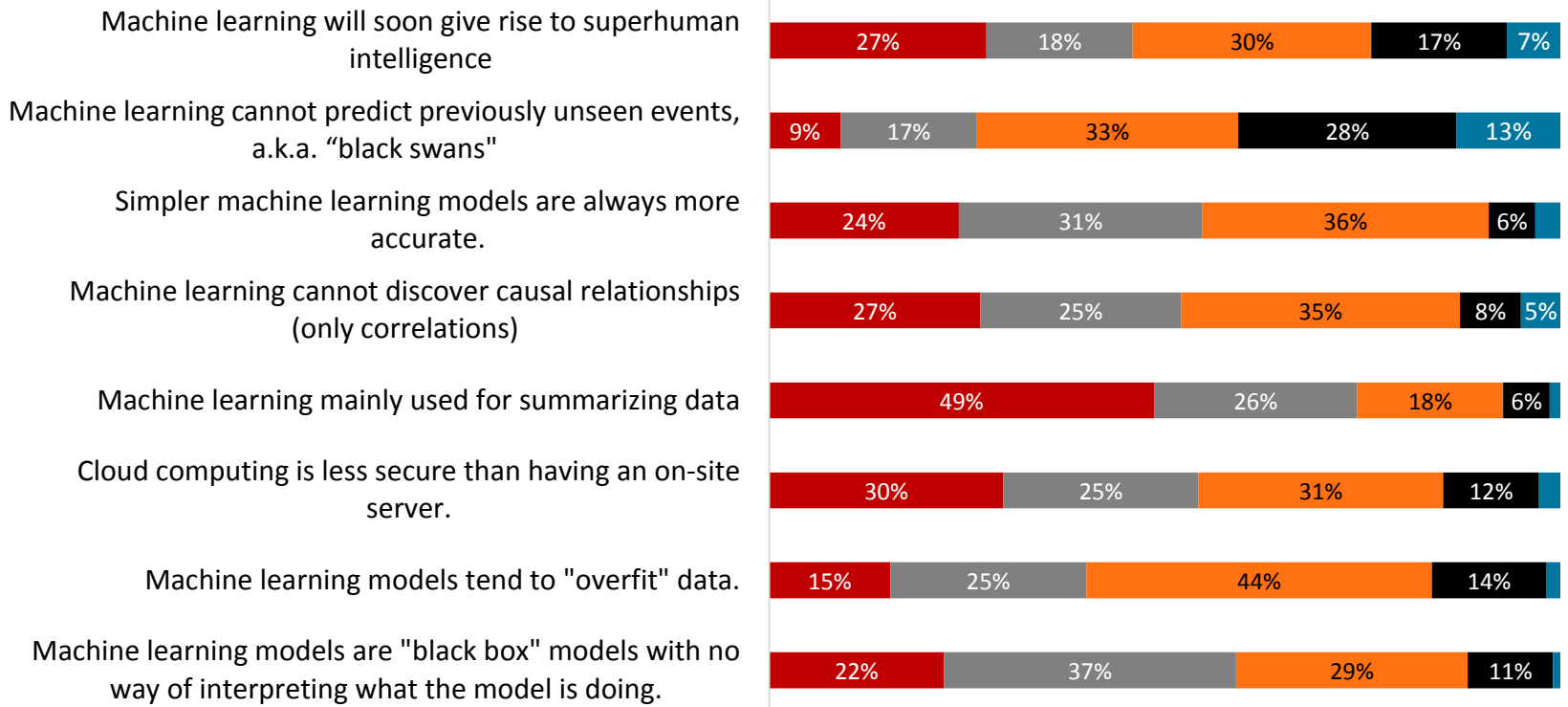


# Member opinions



# Member opinions

## Opinions

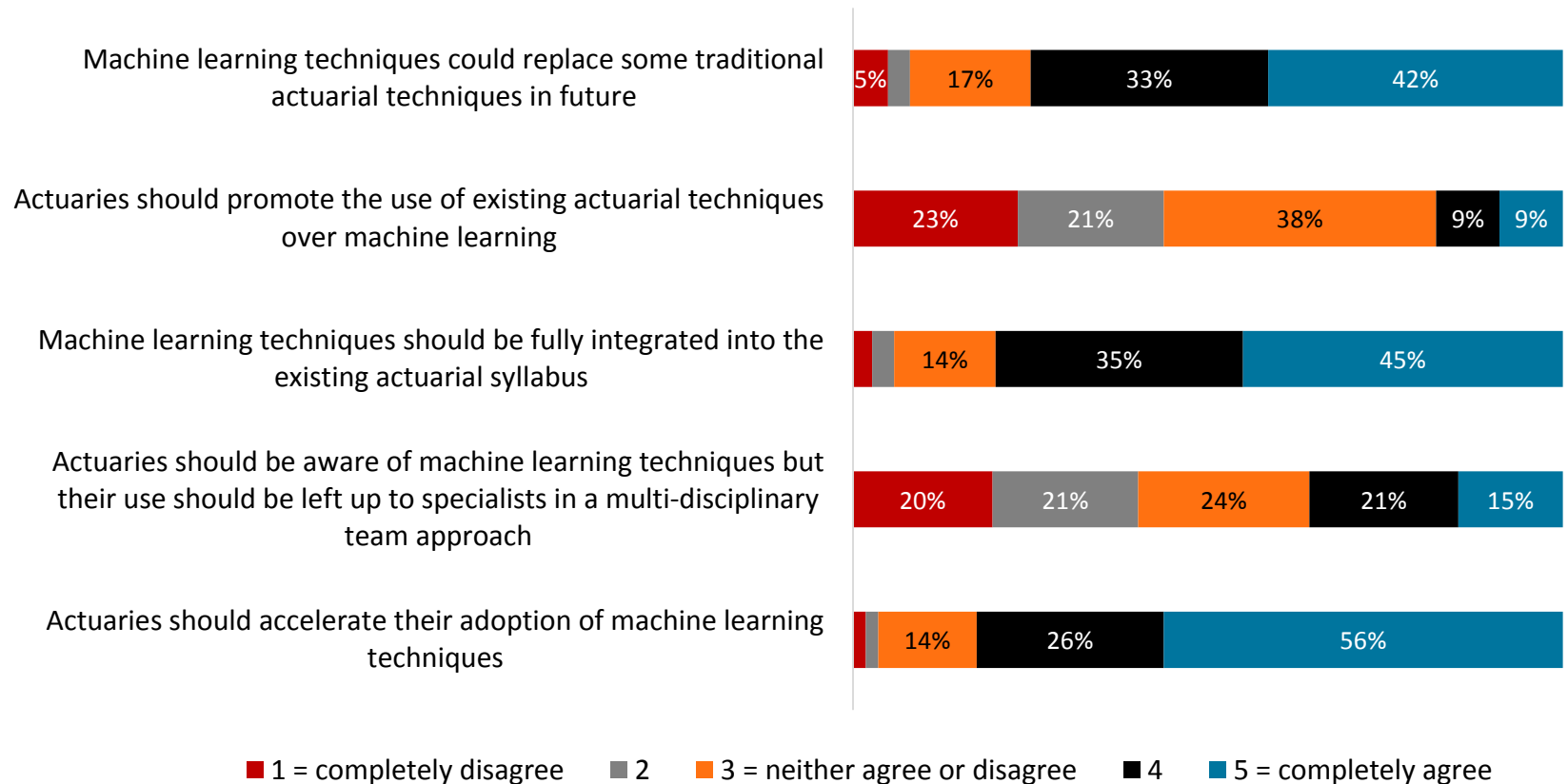


■ 1 = Completely disagree  
 ■ 2  
 ■ 3 = Neither agree or disagree  
 ■ 4  
 ■ 5 = Completely agree



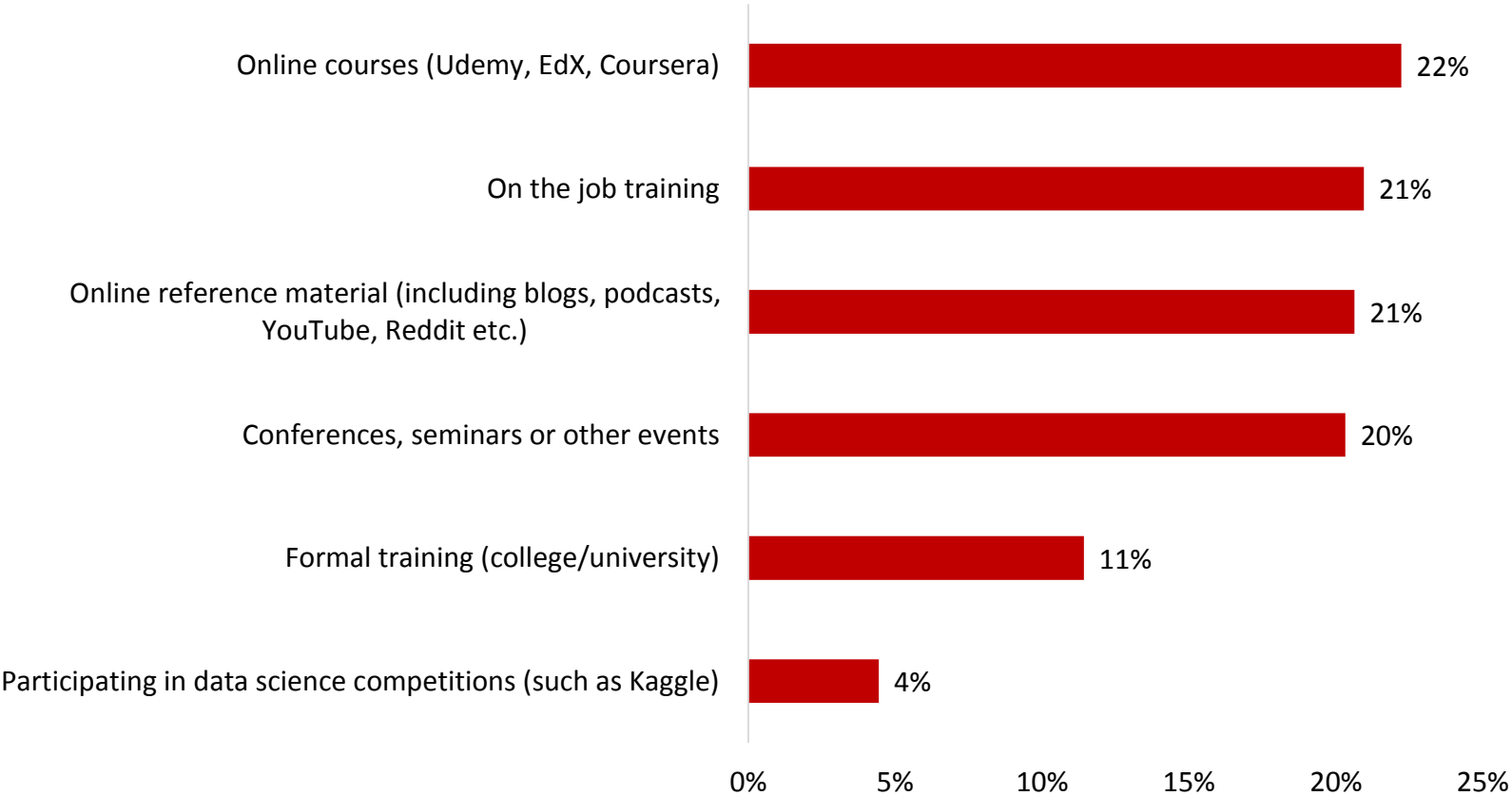
# Machine learning in the actuarial profession

## The place of machine learning in the actuarial profession



# Source of Data Science knowledge acquisition

Source of Data Science knowledge acquisition



# Agenda

Members

Organizations

Tools and  
Techniques

Risks and  
opportunities

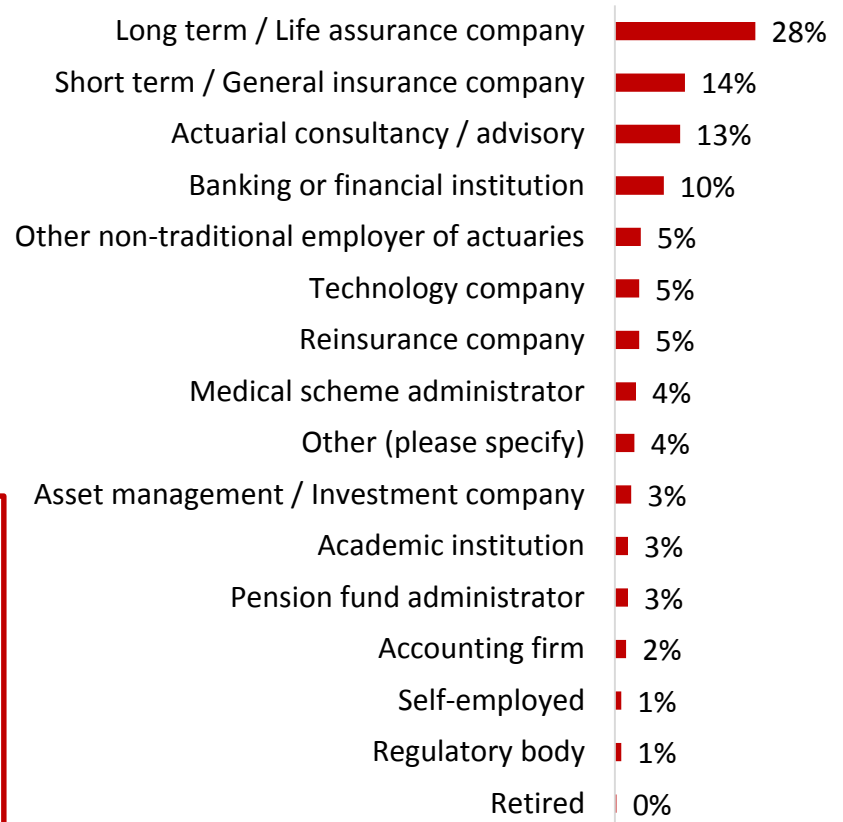
The survey

# Organization type

## Service area



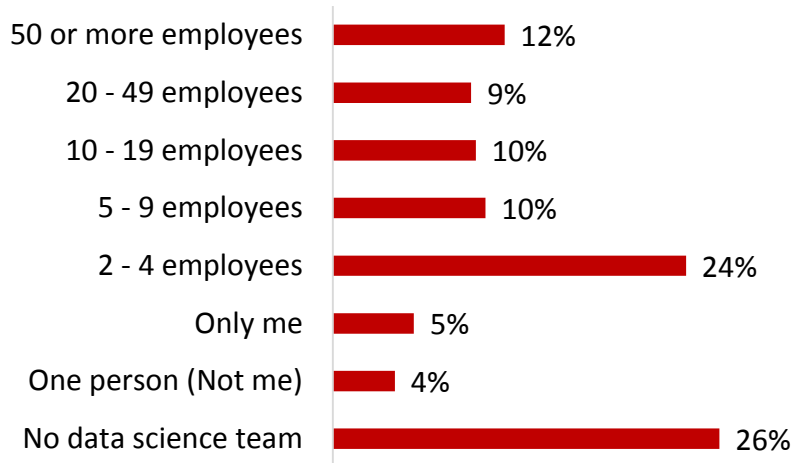
## Organisation type



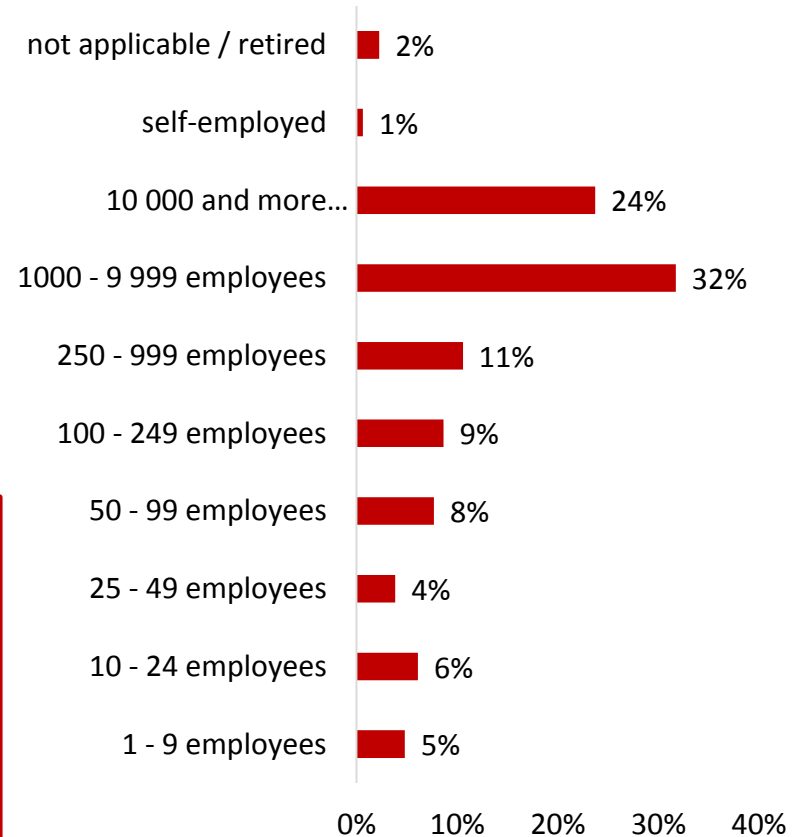
The majority of respondents are employed by traditional actuarial employers, but a fair portion of the respondents are applying their skills in non-traditional areas.

# Size

## Data science team size



## Size of organisation

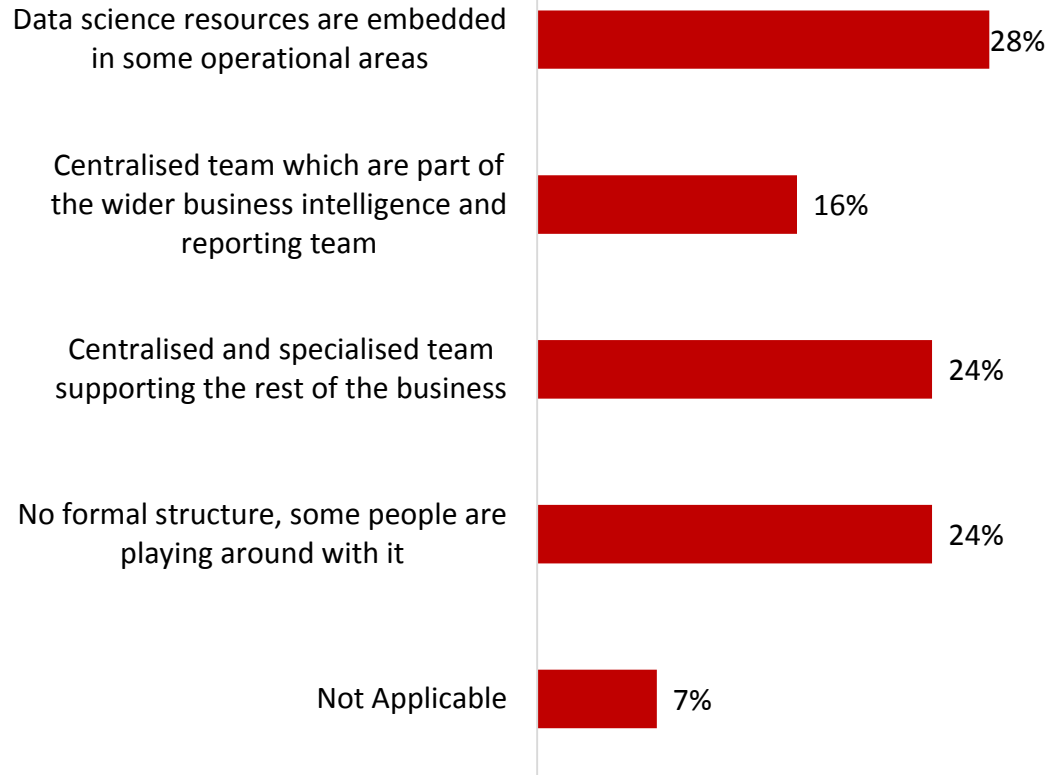


- Most respondents work for larger organisations.
- Many organisations have no data science teams.
- Some seem to be starting the data science journey with teams between 2 and 4 while others have larger, more mature data science teams.

# Centralisation of data science function

## Centralisation approach

The decision around centralization seems to be diverse across organisations.



# Agenda

Members

Organizations

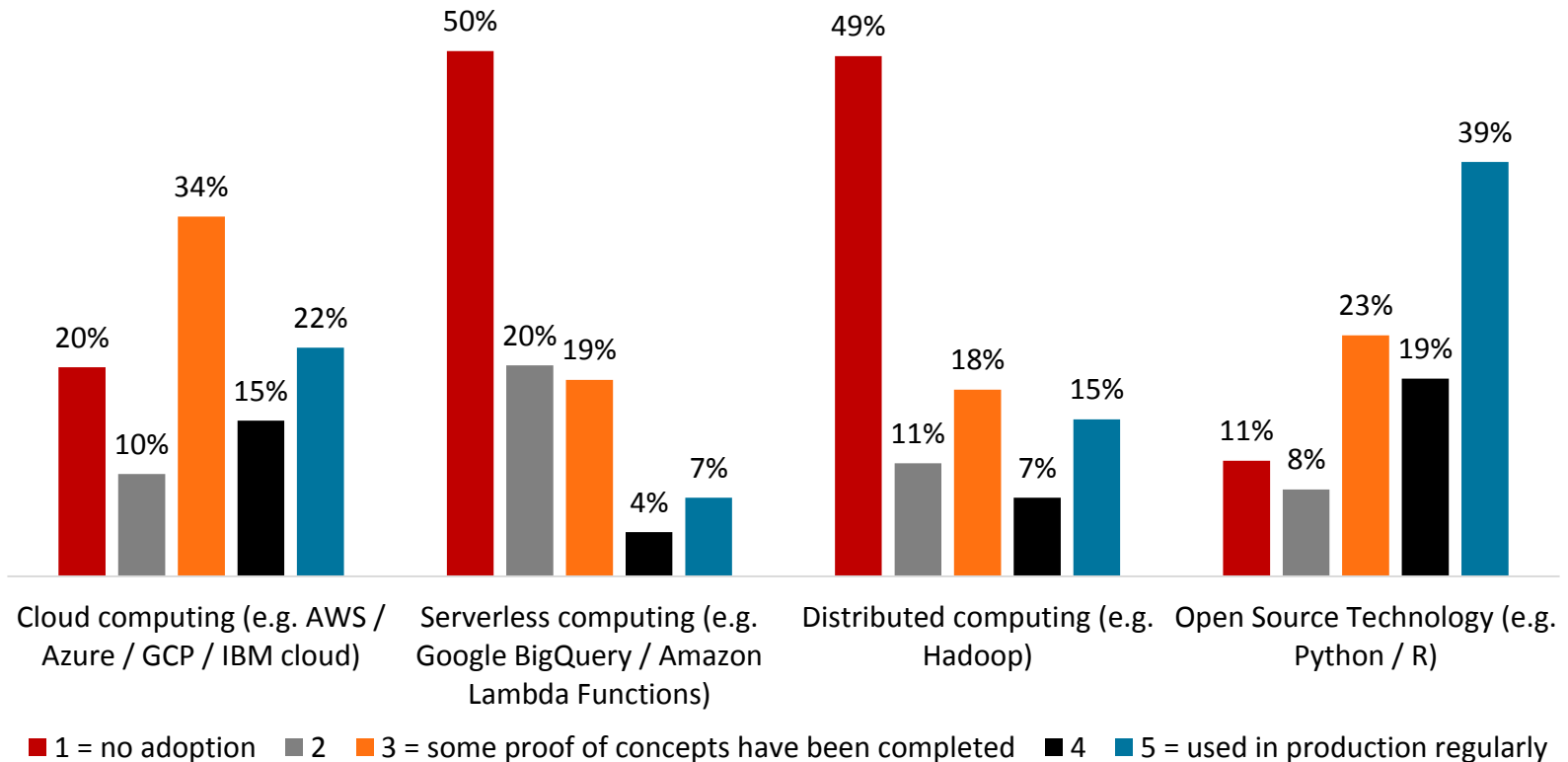
Tools and  
Techniques

Risks and  
opportunities

The survey

# Organizational adoption

## Organisational adoption

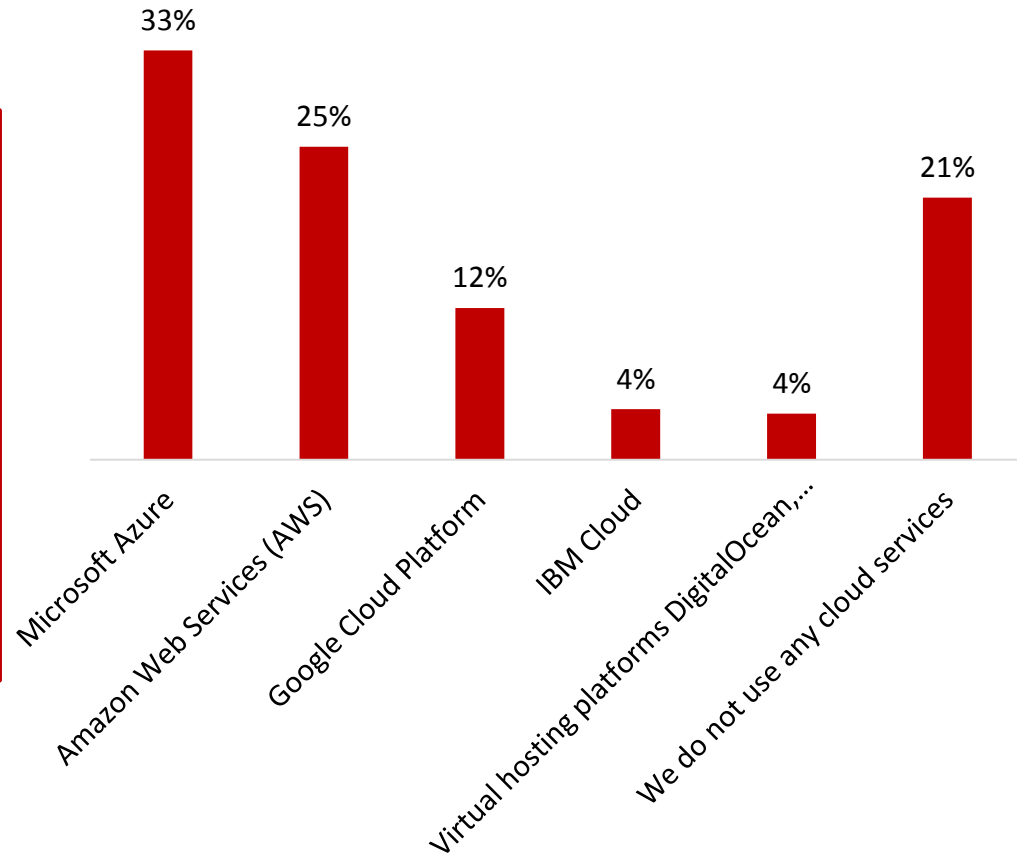




# Organizational adoption

## Cloud provider adoption

Relatively high representation of other cloud providers compared to Amazon Web Services in the context of international market share.



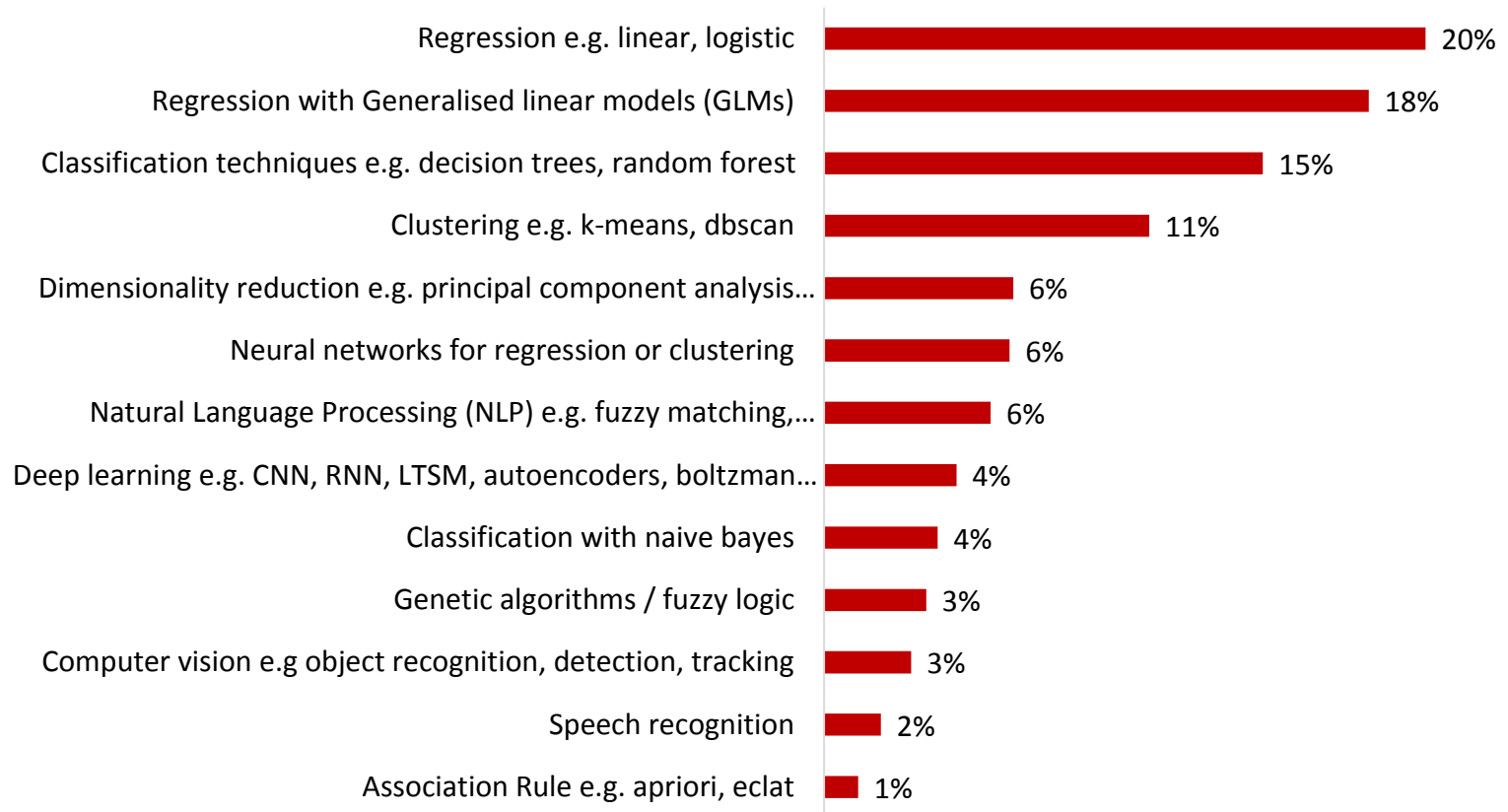
# Machine learning techniques



This sums up my machine learning experience.

# Machine learning techniques

## Machine learning technique



# Coding

| Language                               | 1 - Never heard of it | 2 - Heard of it | 3 - Can read it | 4 - Can write code | 5 - Proficient |
|--|-----------------------|-----------------|-----------------|--------------------|----------------|
| SQL                                    | 1%                    | 14%             | 11%             | 33%                | 40%            |
| Visual Basic                           | 1%                    | 17%             | 24%             | 36%                | 21%            |
| R                                      | 1%                    | 23%             | 19%             | 34%                | 23%            |
| Python                                 | 0%                    | 36%             | 17%             | 28%                | 19%            |
| SAS                                    | 3%                    | 42%             | 17%             | 21%                | 17%            |
| C/C++                                  | 3%                    | 48%             | 27%             | 19%                | 3%             |
| Matlab                                 | 5%                    | 47%             | 21%             | 24%                | 3%             |
| C#                                     | 7%                    | 51%             | 26%             | 12%                | 5%             |
| Java                                   | 0%                    | 69%             | 21%             | 8%                 | 2%             |
| Javascript / Typescript / Coffeescript | 6%                    | 69%             | 18%             | 6%                 | 1%             |
| PHP                                    | 48%                   | 43%             | 6%              | 3%                 | 0%             |
| Julia                                  | 56%                   | 38%             | 3%              | 3%                 | 0%             |
| Golang                                 | 75%                   | 23%             | 1%              | 1%                 | 0%             |
| Rust                                   | 73%                   | 26%             | 1%              | 0%                 | 0%             |

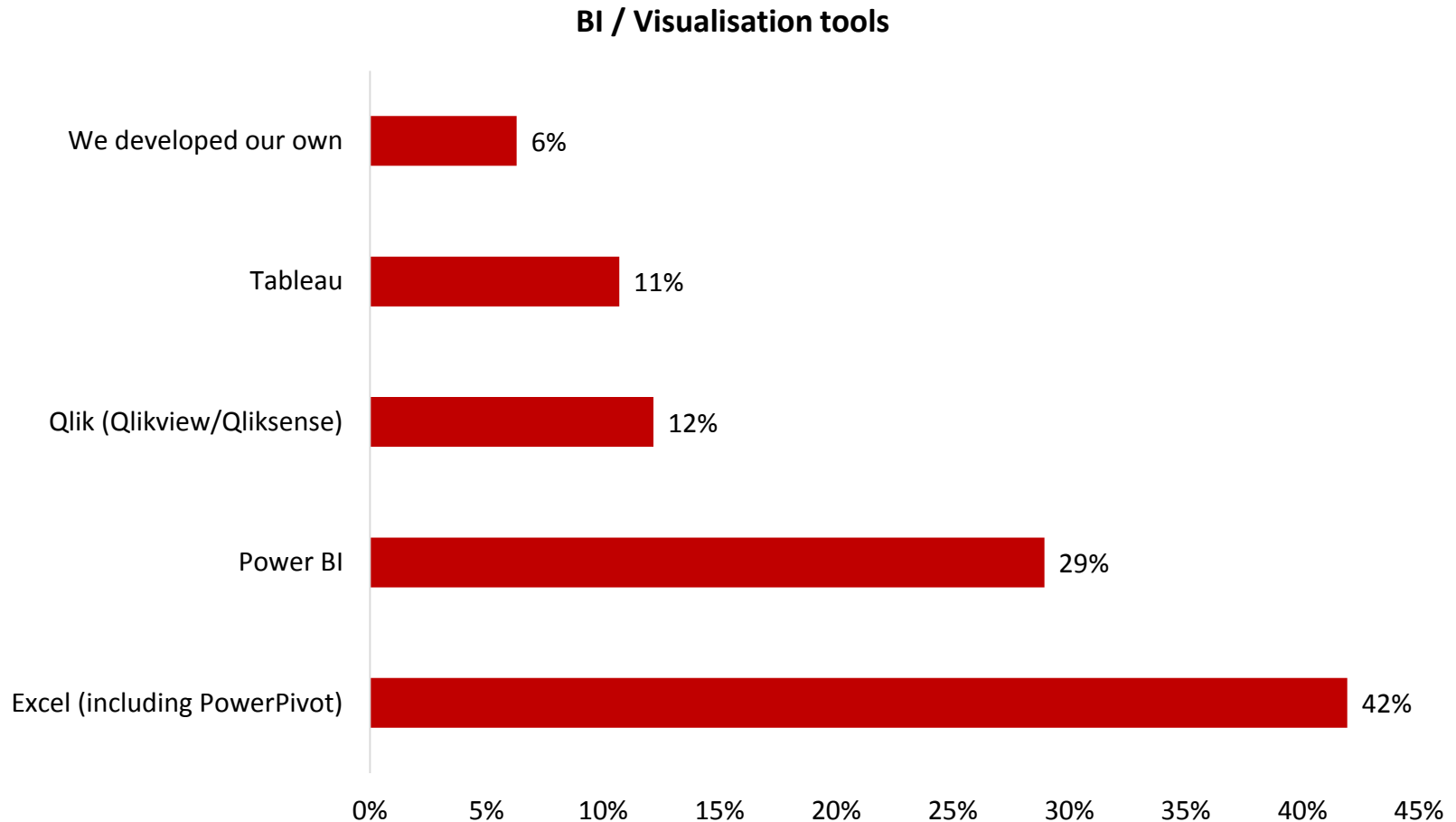
# Coding

| Language     | 1 - Never heard of it | 2 - Heard of it | 3 - Can read it | 4 - Can write code | 5 - Proficient |
|--------------|-----------------------|-----------------|-----------------|--------------------|----------------|
| SQL          | 1%                    | 14%             | 11%             | 33%                | 40%            |
| Visual Basic | 1%                    | 17%             | 24%             | 36%                | 21%            |
| R            | 1%                    | 23%             | 19%             | 34%                | 23%            |
| Python       | 0%                    | 36%             | 17%             | 28%                | 19%            |
| SAS          | 3%                    | 42%             | 17%             | 21%                | 17%            |
| Julia        | 2%                    | 18%             | 27%             | 18%                | 35%            |

The relatively high adoption of R compared to Python is interesting, especially as the 2019 Kaggle ML and Data Science Survey had Python at 87% and R at 31%.

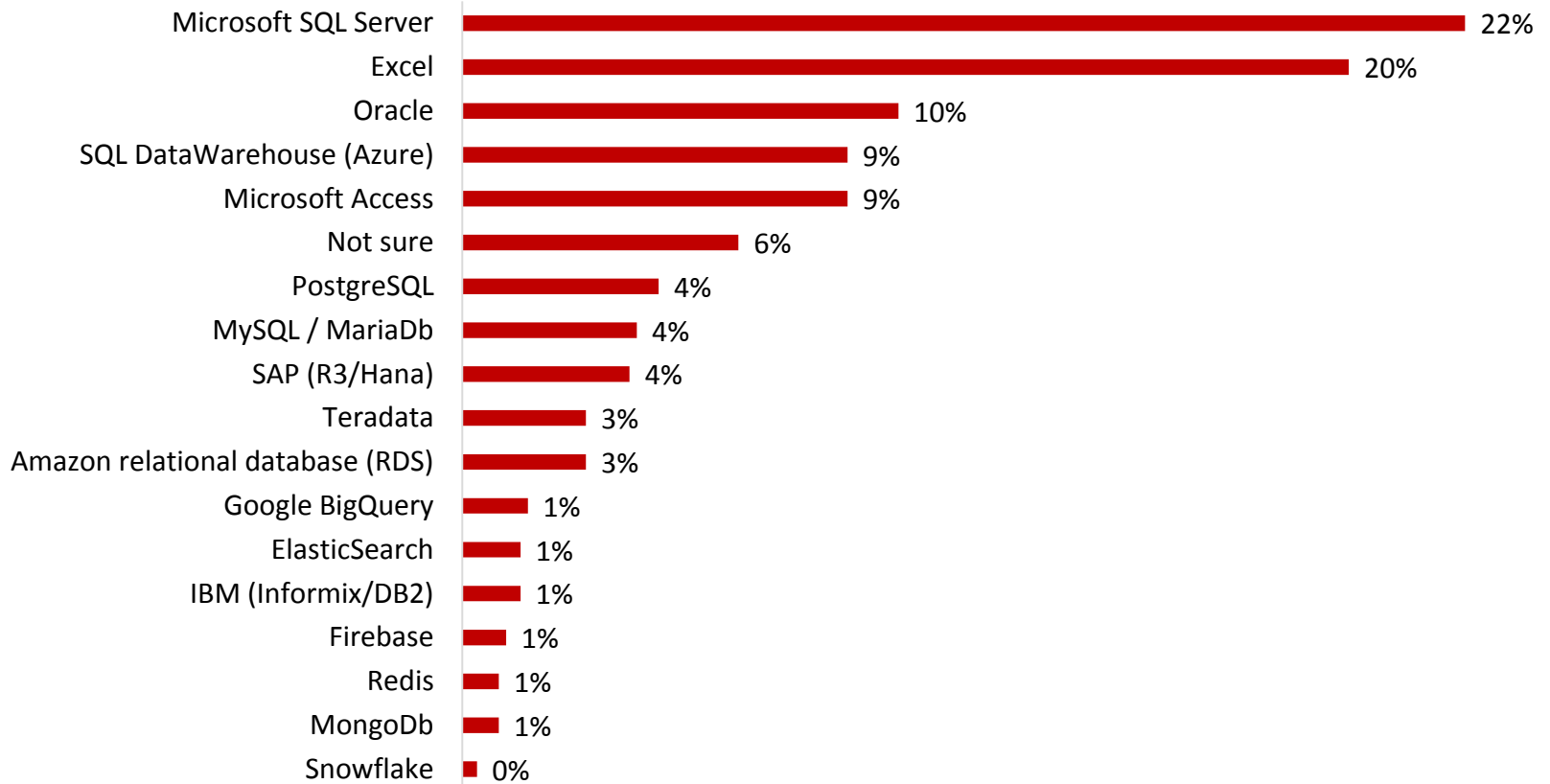
The high adoption of Visual Basic is to be expected, considering the high usage of Excel in the actuarial space.

# BI / Visualisation tools



# Data Storage System

Data Storage System



# Agenda

Members

Organizations

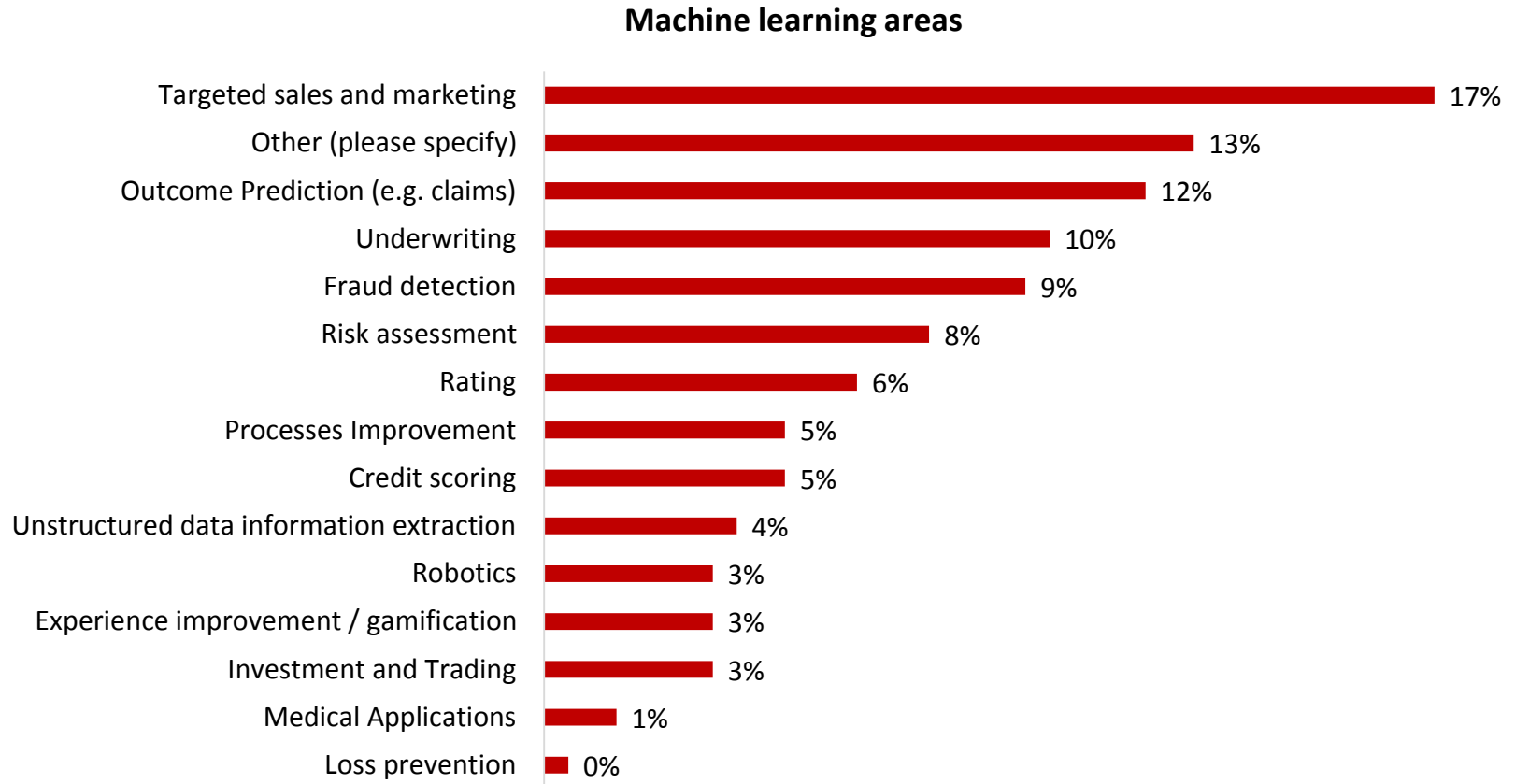
Tools and  
Techniques

Risks and  
opportunities

The survey



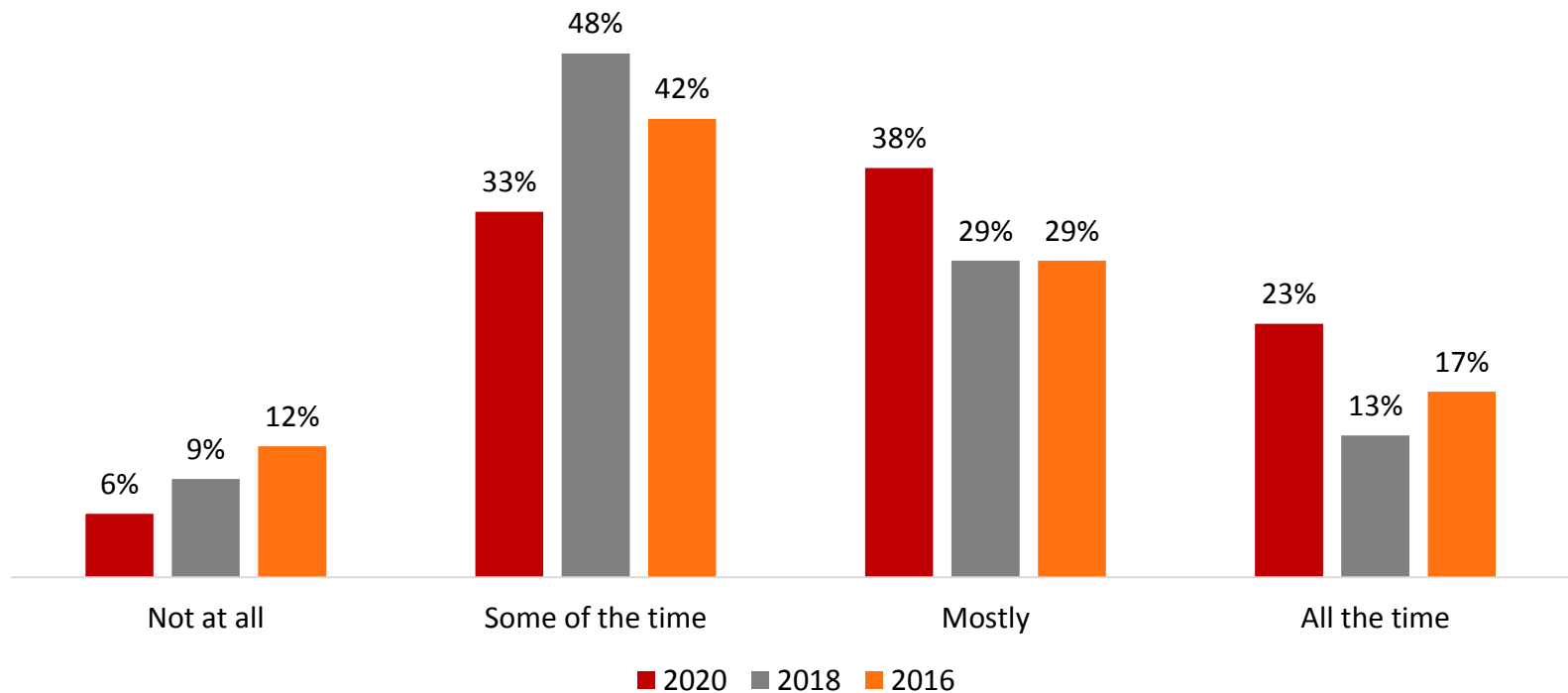
# Machine learning areas



\* Note: Due to a technical issue, only one area was selectable in the survey

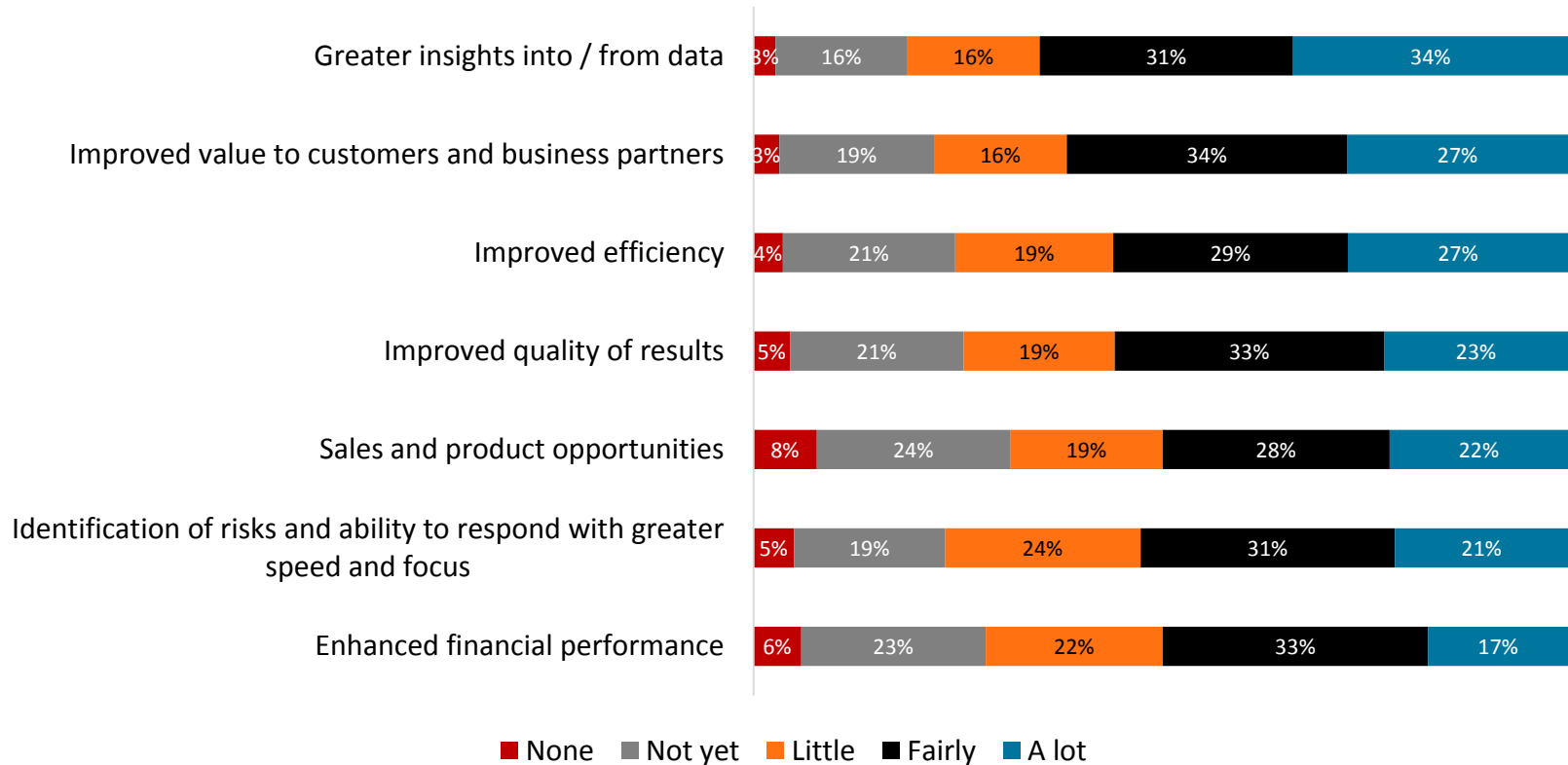
# Decision making

How often are important decisions made in your company informed and supported by data analytics? (or clients if you are a consultant)



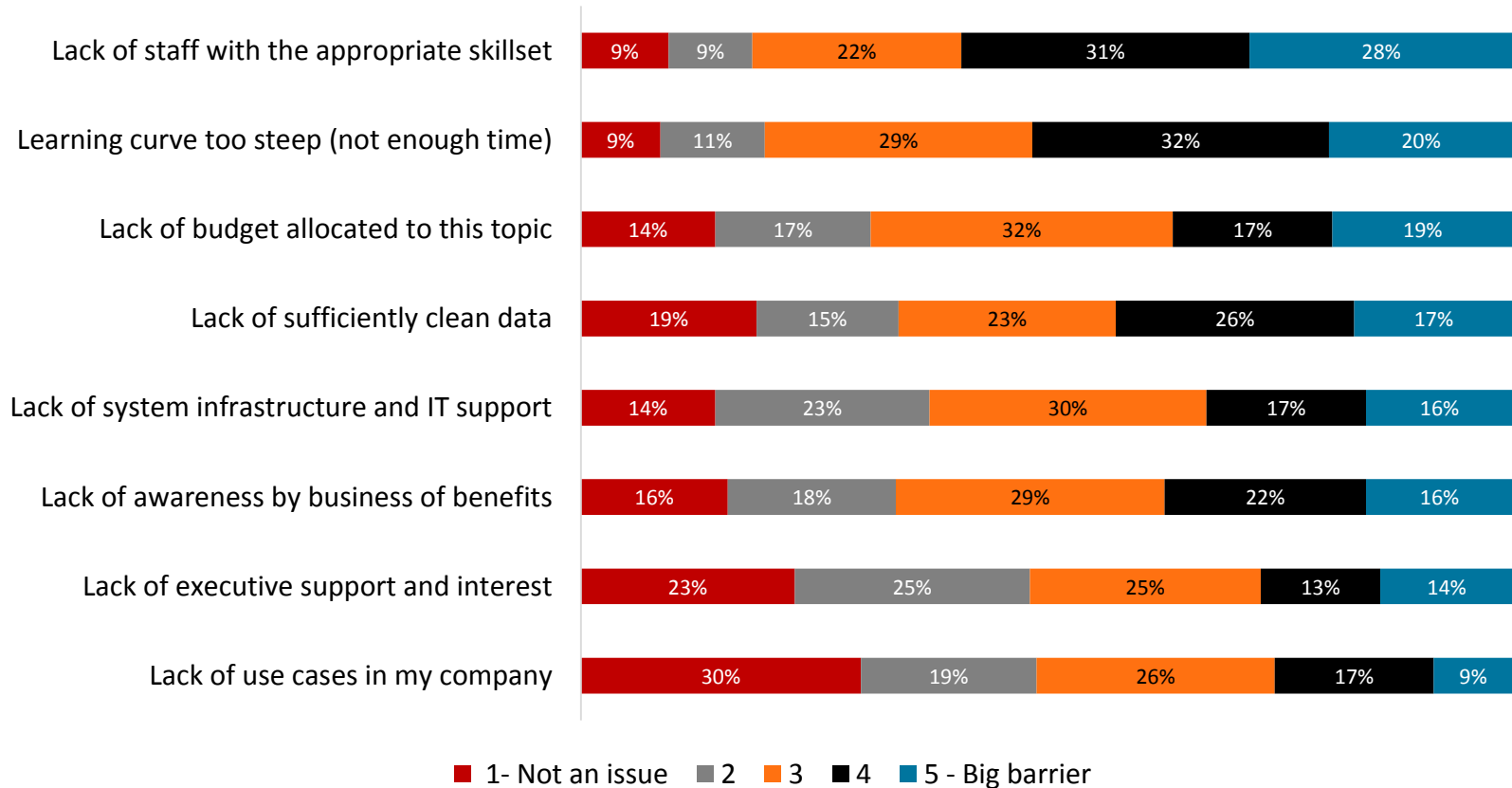
# Data science benefits

Rate to what extent (if any) you think data science and machine learning have benefited your company



# Data science barriers

## Barriers to progress



# Agenda

Members

Organizations

Tools and  
Techniques

Risks and  
opportunities

The survey

# Future improvements to the survey

- Fewer questions
- Fewer multiple choice options
- Options to express opinions in free text
  - Bugs in the survey
  - Include unsure options
  - Release response data?
- Focus on other areas of data science

Thank you...

Questions